DOCUMENT RESUME

ED 254 238 IR 051 040

AUTHOR de Man, W. H. Erik, Ed.
TITLE Conceptual Framework and Guidelines for Establishing Geographic Information Systems Capable of Integrating Natural Resources Data and Socio-Economic Data for Development-Oriented Planning, Monitoring and Research. (Preliminary Version).
INSTITUTION United Nations Educational, Scientific and Cultural Organization, Paris (France). General Information Programme.
REPORT NO PGI-84/WS/20
PUB DATE Aug 84
NOTE 74p.
PUB TYPE Guides - Non-Classroom Use (055) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Cooperative Planning; *Databases; *Developing Nations; Geographic Distribution; Guidelines; *Information Networks; *Information Systems; Information Utilization; *Natural Resources; Physical Geography; Regional Planning; Scientific and Technical Information; Socioeconomic Influences
IDENTIFIERS *Geobased Information Systems; Geocoding; *Spatial Planning

ABSTRACT

This preliminary conceptual framework and the 23 related guidelines have been prepared by specialists drawn from both developed and developing nations as a joint project of the International Institute for Aerial Survey and Earth Sciences (ITC). The project was designed to analyze linkages between natural resource based data and certain types of socioeconomic, scientific, and technological information, and to prepare guidelines to serve as an operational framework for establishing integrated spatial data banks, with particular emphasis on ensuring their usefulness in development planning. The report deals with geographic information systems from three perspectives: as information systems, having particular importance in the development process; in terms of the specific spatial function they perform; and the challenges presented by the integration of spatial data. Following an introduction and summary of the guidelines, the report is divided into five sections: (1) the nature of geographic information systems; (2) general considerations relevant to information systems (including identification of the users and uses of integrated spatial information and design considerations); (3) collection and handling of spatial data, including geocoding; (4) establishing an integrated geographic information system; and (5) post-establishment managerial concerns. A list of participants and resource persons contributing to the project, a selective list of existing geographic information systems, and a glossary of terms used in the report are appended. (THC)

ED254238

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

PGI-84/WS/20

Conceptual framework and guidelines for establishing geographic information systems

capable of integrating natural resources data and socio-economic data for development-oriented planning, monitoring and research

(Preliminary version)

Reproduced by permission of
UNESCO

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY
C. Coudert Schklowski
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

General Information Programme and UNISIST

United Nations Educational
Scientific and Cultural Organization

Paris
August 1984

IR 051040
IR051040

PGI-84/WS/20
Paris, August 1984

CONCEPTUAL FRAMEWORK
AND
GUIDELINES
FOR ESTABLISHING GEOGRAPHIC INFORMATION SYSTEMS

capable of integrating natural resources data and
socio-economic data for development-oriented planning,
monitoring and research

(Preliminary version)

Edited by W. H. Erik de Man

General Information Programme and UNISIST
United Nations Educational,
Scientific and Cultural Organization

The designations employed and the presentation of material in this work do not imply the expression of any opinion whatsoever on the part of the Unesco Secretariat concerning the legal status of any country or territory, or of its authorities, or concerning the delimitations of the frontiers of any country of territory.

Recommended catalogue entry :

de Man, W. H. Erik
Conceptual framework and guidelines for establishing geographic information systems capable of integrating natural resources data and socio-economic data for development-oriented planning, monitoring and research : preliminary version / edited by W. H. Erik de Man [for the] General Information Programme and UNISIST .- Paris : Unesco, 1984 .-vi, 67 p., 30 cm .- (PGI-84/WS/20)

I - Title
II - Unesco. General Information Programme and UNISIST

© Unesco, 1984

BEST COPY AVAILABLE

PREAMBLE

The present preliminary conceptual framework and guidelines have been prepared by specialists drawn from both developed and developing nations as a joint project of the International Institute for Aerial Survey and Earth Sciences (ITC), Enschede, The Netherlands, and UNESCO within its General Information Programme. The various activities of the project were co-ordinated by the ITC.

The terms of reference of the project were "to analyze linkages between natural resource based data and certain types of socio-economic, scientific and technological information and to prepare guidelines to serve as an operational framework for establishing integrated spatial data banks, with particular emphasis on ensuring their usefulness in development planning".

This project was the result of a recommendation by the Meeting of Experts on Provision of Data for Development, UNESCO, 24-27 November, 1981, and was completed in June, 1984. The terms of reference of this project were shaped by the recommendations made at a preceding Workshop on Information Requirements for Development Planning in Developing Countries, held at ITC, Enschede, The Netherlands, February 18-23, 1980.

Through a number of rounds of correspondence and subsequent comments on successive drafts of the guidelines, the co-ordinators collected contributions and experiences from the field. The report was finalised at an Experts Meeting held at ITC from 27 February to 2 March, 1984.

The assistance of the worldwide group of individuals and institutions who have contributed and actively participated in this project is gratefully acknowledged. The participants of the Experts' Meeting and the group of collaborators are listed in Annex 1.

It is hoped that the preliminary conceptual framework and guidelines will be applied in a series of pilot projects under varying conditions and that the experiences so gained can be used to refine the approach in the future. Comments and suggestions from the readership are most welcome; they should be addressed to:

Division of the General Information Programme
UNESCO
7, Place de Fontenoy
75700 PARIS
France

TABLE OF CONTENTS

		Page
0.	Introduction	1
	Summary of the guidelines	3
1.	The nature of geographic information systems	10
	1.1 Information systems	10
	1.2 Spatial data and information	12
	1.3 Integration of data	13
2.	General considerations relevant to information systems	14
	2.1 Identification of the users and uses of integrated spatial information	15
	2.1.1 Orientation of an information system	15
	2.1.2 Identification of the users	16
	2.1.3 Utilization of information	18
	2.1.3.1 Research	19
	2.1.3.2 Inventories and design studies	20
	2.1.3.3 Monitoring and evaluation	20
	2.1.3.4 Underpinning of 'informal information'	22
	2.2 Design considerations of information systems	23
	2.2.1 Existing information capabilities and practices	23
	2.2.2 Hierarchical systems and decentralized information systems	24
	2.2.3 Skills	26
	2.2.4 Technology	27
	2.2.5 Change strategy	28
3.	Collection and handling of spatial data; geocoding	30
	3.1 Alternative methods of geocoding	30
	3.2 Application of different methods of geocoding	34
	3.3 Choice of the appropriate method of geocoding; trade-offs	36
	3.3.1 Spatial variability	37
	3.3.2 Possibilities and limitations in data collection	38
	3.3.3 Context of decision-making	39

Page

4. Establishing an integrated geographic information system 41

4.1 Integration of different sets of spatial data 41

4.2 Co-ordination and institutional framework 44

5. Post-establishment managerial concerns 46

5.1 Technical support 46

5.2 Financial support 47

5.3 Logistic support 47

5.4 Management of change 48

References cited 49

Annex 1 List of participants of Experts' Meeting and list of resource persons, contributing to the project 52

Annex 2 Selective list of existing geographic information systems 59

Annex 3 Glossary of terms used in the report 65

BEST COPY AVAILABLE

0. INTRODUCTION

The conceptual framework and the guidelines presented in this report are addressed to development programme administrators, planning specialists, resource managers and policy makers with the aim of facilitating communication between the above and those who will implement and operate the geographic information systems. This report deals with geographic information systems from three perspectives:

1. as information systems, having particular importance in the development process,
2. in terms of the specific spatial function they perform, and
3. the challenges presented by the integration of spatial data.

While geographic information systems are placed within the general context of information systems (e.g. concerning their institutional setting and support) emphasis will be placed on their spatial data processing functions and especially those functions which deal with the integration of natural resources data and socio-economic data.

There is no universal approach to the development of capabilities in handling and provision of spatial information. Moreover, the level of available and/or possible technology may differ from one country to another. For example it may not be feasible to implement a very advanced system in a short time period in a given national situation and, in this case, a phased implementation is indicated and alternative routes should be considered.

In this context, the present guidelines are not intended as a 'blue-print' for information systems capable of integrating different sets of spatial data, but aim rather to contribute to and have an impact on the on-going process of development, the first step of which is the creation of a minimum degree of awareness. Nevertheless, it is hoped that the document is sufficiently detailed to be of use in a wide variety of development situations.

For the benefit of the reader a summary of the guidelines is presented below, preceded by a flow diagram indicating their inter-relationships and stages of application.

BEST COPY AVAILABLE

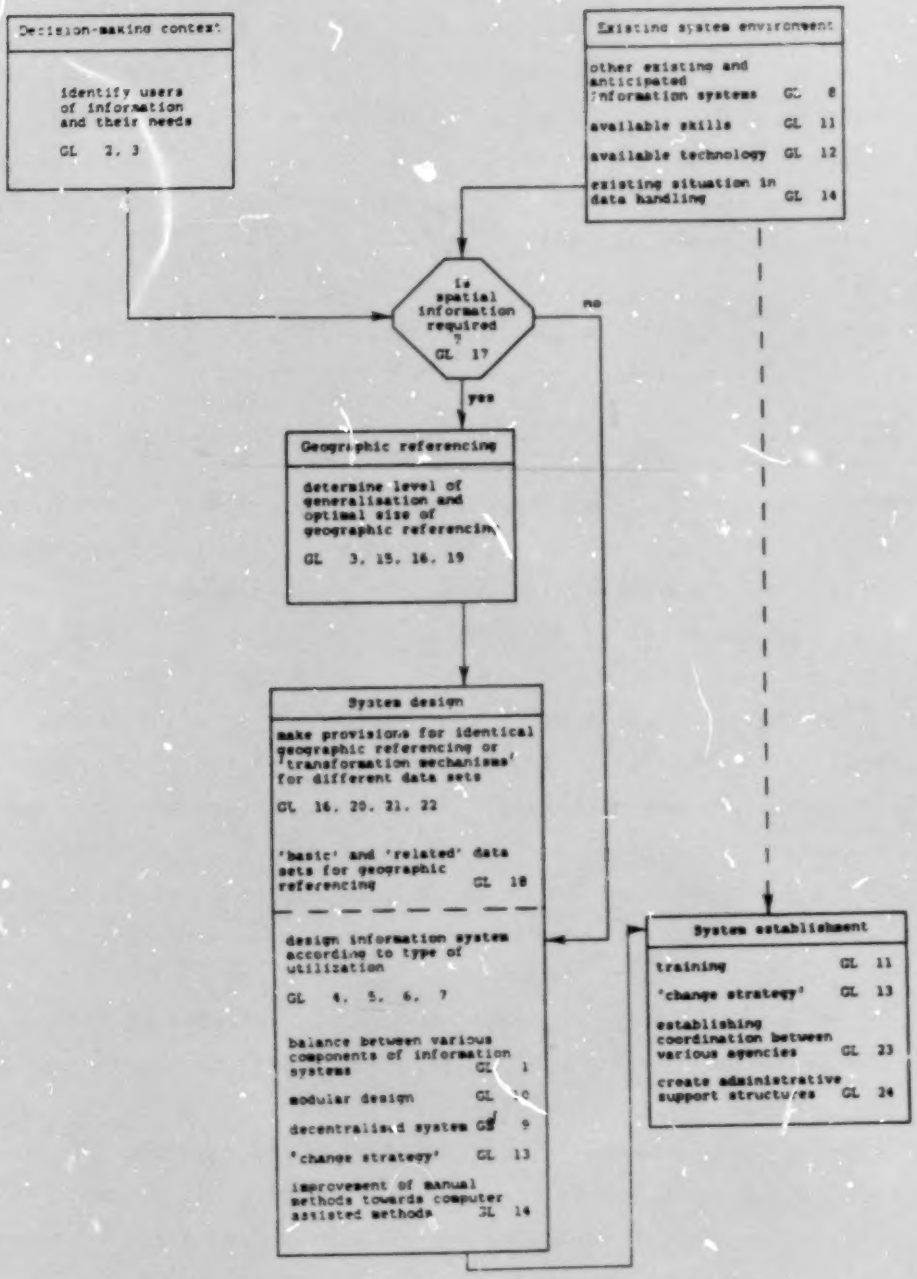

Decision-making context
identify users of information and their needs
GL 2, 3
Existing system environment
other existing and anticipated information systems GL 8
available skills GL 11
available technology GL 12
existing situation in data handling GL 14
Is spatial information required ? GL 17
no
yes
Geographic referencing
determine level of generalisation and optimal size of geographic referencing
GL 3, 15, 16, 19
System design
make provisions for identical geographic referencing or 'transformation mechanisms' for different data sets
GL 16, 20, 21, 22
'basic' and 'related' data sets for geographic referencing GL 18
design information system according to type of utilization
GL 4, 5, 6, 7
balance between various components of information systems GL 1
modular design GL 10
decentralised system GL 9
'change strategy' GL 13
improvement of manual methods towards computer assisted methods GL 14
System establishment
training GL 11
'change strategy' GL 13
establishing coordination between various agencies GL 23
create administrative support structures GL 24

Fig. 1: Flow diagram showing relations among the guidelines (the numbers following the letters 'GL' refer to guidelines developed in this document).

SUMMARY OF GUIDELINES
FOR ESTABLISHING GEOGRAPHIC INFORMATION SYSTEMS
capable of integrating natural resources data and
socio-economic data for development-oriented planning,
monitoring and research

GUIDELINE 1: A balanced development of the information system an of its components (conceptual base; data collection, storage and processing; as well as analytical capabilities) should be aimed at.

GUIDELINE 2: In order to identify the (potential) users of information, it is suggested that inventories be made of:
- existing flows and utilization of data/information,
- the relevant decision processes and potential information users at different levels of responsibility.

These inventories can be confined to the specific sectors and/or user communities of the information system in question. The emphasis in approach will depend on the orientation of the information system, i.e. general purpose, supply-oriented or function-oriented.

The execution of a pilot survey in collaboration with the user community may facilitate a subsequent full inventory of the needs of users and meaningful user participation in the envisaged system.

GUIDELINE 3: Before considering the design of an information system, the types of utilization situations (e.g. research, inventories, monitoring/evaluation, 'underpinning of informal information') and the required degree of accuracy, precision and resolution in data must be identified.

GUIDELINE 4: Information systems serving a heterogeneous group of clients/users who face different situations should be flexible and adaptable to permit easy inclusion and retrieval of detailed information when needed. This generally requires some standardization of concepts and formats for data collection.

GUIDELINE 5: Information systems for monitoring or evaluation of specific projects and programmes should be embedded in the design of these projects and programmes, whereas more general information systems for monitoring or evaluation should be embedded in the management and decision-making structures of the information users.

GUIDELINE 6: Monitoring requires the establishment of sets of key-indicators and an information system that can adequately cope with them.

GUIDELINE 7: When designing an information system, the availability of 'informal information' should not be overlooked; indeed, such information should be utilized to the maximum extent possible.

GUIDELINE 8: The design and creation of an information system should be placed in the context of other information systems, existing as well as anticipated. To this end, a survey of existing data and of information-handling capabilities and practices should be made before embarking on the design of a new system.

GUIDELINE 9: Decentralized information systems in general have a great potential for facilitating local and sectoral decision-making. Inexpensive data processing technology is becoming available to implement such systems. Decentralization, however, requires careful formulation and regulation of the linkages between the central unit and the peripheral nodes as well as a strengthening of local capabilities.

GUIDELINE 10: At an early stage in the design of an information system, training programmes have to be developed to guarantee the availability of skills. The resources needed for these training programmes must be ensured.

GUIDELINE 11: The envisaged information system should take into account both locally available technology (including hardware and software as well as manpower) and international trends in technological development.

GUIDELINE 12: In parallel with the selection of a computer system, a careful strategy for change and adaptation should be developed. Data collection procedures should depend as little as possible on the system's hardware in order to facilitate necessary future changes.

GUIDELINE 13: It is generally preferable that information systems develop within existing administrative and institutional contexts. It is usually advisable to examine the evolution of manual methods into computer assisted methods. These improvements should show a benefit as soon as possible in order to sustain the confidence and interest of the users.

GUIDELINE 14: The optimal size of the basic spatial units of an individual data set depends on the required spatial precision of the phenomena considered. Relatively large basic spatial units are often acceptable, thus simplifying the sampling procedures needed to collect the appropriate data.

GUIDELINE 15: The optimal level of spatial detail for a geographic information system as a whole represents a compromise with respect to the requirements for each of the individual data sets.
As a means of improving processing and storage efficiency, individual data sets with exceptional degrees of detail can be regarded as separate but co-ordinated data sets.

GUIDELINE 16: An appropriate method of geographic referencing can only be selected after determining the requirements in terms of treatment of spatial data, the required level of spatial detail and the need to integrate specific sets of spatial information.

GUIDELINE 17: Whenever practical, fully compatible geocoding schemes should be adopted at the stage of data collection. Otherwise, methods have to be found to transform or adjust the geocoding schemes applied to individual data sets to permit matching with other data sets as necessary. This question should be considered with care since in specific situations mixed geocoding schemes can be more cost-effective than having a single scheme.

GUIDELINE 18: An integrated geographic information system may be composed of separate but co-ordinated data sets each having a different method of geographic referencing. Of these data sets, one or more are selected as the basic data set(s), while the others are considered as related data sets. The basic data sets have a stronger influence on the methodology for integrating the various data sets.

GUIDELINE 19: Certain kinds of data sets on natural resources which have a high degree of variation over space and a low decay rate are appropriate basic data sets within integrated information systems.

GUIDELINE 20: When integrating spatial data sets which differ in terms of scale and/or detail, care should be taken to inform the users of the inherent limitations of the data.

GUIDELINE 21: The integrated geographic information system should be designed in a modular way. It may prove useful to foresee a common 'data transformation service' which may be either manual or automated.

GUIDELINE 22: An inter-agency or interdepartmental steering committee is an effective instrument to plan, monitor and evaluate the development and use of an integrated geographic information system. In such a steering committee both users and producers of spatial data should be represented. The committee could be housed within an appropriate government policy or planning agency.

GUIDELINE 23: To ensure smooth and a continuous operation of the information system, administrative support structures have to be created in the technical, financial, logistical and managerial areas. These support structures have to be carefully planned.

BEST COPY AVAILABLE

1. THE NATURE OF GEOGRAPHIC INFORMATION SYSTEMS

1.1 INFORMATION SYSTEMS

Geographic information systems serve as providers of spatial information as defined in the next section. In many respects geographic information systems are similar to non-geographic ones. They share the abilities to accept, process and present data, to update and modify data and to combine data sets orginating from different sources, as shown schematically in Figure 2.

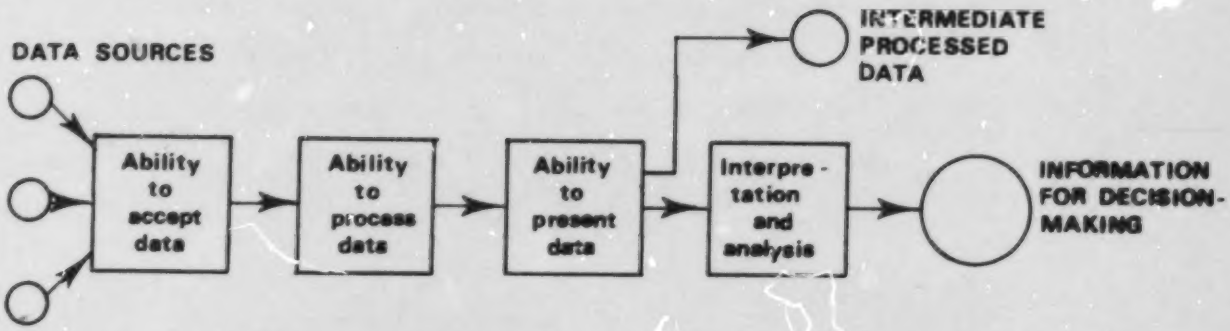


Fig. 2: Sequence of functions of an information system.

BEST COPY AVAILABLE

Many of the tools developed for non-geographic information systems are applicable to geographic information systems. However, the handling of spatial data typically requires specialized and complex tools for data capture, storage, retrieval and presentation. In addition, the multi-source origin of spatial data challenges our ability to integrate them for effective use.

In recent years computers have become the backbone of information systems in the industrialized societies and, also for some applications, in less developed nations. Digital data handling enables easy access and retrieval from large data sets and allows for calibration of data, linkages between data sets and use of new presentation methods.

Computer-based methods are particularly relevant for spatial data. In practice they offer unique facilities for handling the large data sets often associated with spatial data and the ability to combine many different data sets.

An information system may consist of one or more sub-systems. There may be many reasons to design an information system with several sub-systems, e.g. organizational convenience or division of work according to functions (e.g. data collection, data processing) or sectoral activities (agriculture, transportation, mining, etc.). When the various subsystems have standaridized interconnections (and are therefore compatible) they are called 'modules' and the system is called modular. If the sub-systems are at different physical locations then the system is called distributed. The successful implementation of modular and/or distributed systems requires solutions to problems typical of a sub-system environment. These problems include data collection, consistency control, update procedures, failure recovery, and linkage procedures. If each sub-system is allowed to develop in isolation, the ability to integrate information will be limited. Therefore, very clear procedures for interaction and cooperation of the sub-systems have to be established and applied (1).

In cases where information is required in the context of decision-making and planning, particularly at district or local levels, much of this information is often already available as 'informal information' in the form of ideas, feelings experiences, etc., of the individuals involved. While 'informal information' is highly subjective, difficult to verify objectively and typically non-quantitative, it is a rich resource which often tends to be overlooked when (formal) information systems are developed.

Issues like ownership of, accessibility to, and responsibility for data are important legal issues which also have social and professional implications. The design criteria of the information system must thus relate to the current legal, social and political environment and to its likely evolution.

1.2 SPATIAL DATA AND INFORMATION

Data can in general be considered as having three different components (2):
- theme (or content);
- time;
- space (or location).

From the view point of data manipulation (and particularly data integration) it is useful to consider the spatial component as distinct from time and from theme.

Information or data with a spatial component is referred to in this document as simply 'spatial information'. Inherent to spatial information is a specification of the level of spatial detail i.e. the basic area to which the information is referenced such as a plot, village, region or country. Most statistical offices use different levels of spatial resolution in their operations. For example, many statistics are presented at district level or even at municipality level. However, space as a data-dimension is under-utilized at the moment. The small area statistics (e.g. based on the polling district or postal code) are gaining more and more interest, thus facilitating easier integration of different data sets.

If the area under consideration (country, region, municipality, etc.) is divided into small ('basic') spatial units, it is not always necessary to collect all the data for all these units; a sample will often do. If data on items such as living conditions, practices, behaviour, etc. are considered, it is normally quite sufficient to collect information for a relatively small sample of representative spatial units. In some developing countries, it is not only impractical to collect vast amounts of data, but even if collected, such data could not easily be processed, analysed, reported upon or acted upon. It is important to ensure that facilities exist to use data before they are collected.

Whereas computers are helpful tools in the processing of spatial data, the methodologies discussed in this report relate to opportunities and problems that can be realised in both computerized and non-computerized environments. It is clear, however, that technological developments will eventually extend electronic data processing into organizations and countries in which it is not common at present. It is therefore highly advisable to observe conventions for data representation and data consistency which are associated with computerized systems, even though the current system may have a large manual element.

1.3 INTEGRATION OF DATA

A very important aspect of geographic information systems relates to the ability to integrate data from different sources, thereby providing genuinely new information. There is an underlying assumption in this study that data integration must be considered a prime function of a geographic information system.

In many instances, phenomena can only be described adequately with the help of more than one variable. For example, a plot may be described by location, land use and title (ownership and occupancy), or the condition of the soil by several attributes which refer to a specific location. In this example, the present land use can be compared with soil qualities with the help of the variable "location".

"Integration" of data means relating different phenomena on the basis of a common set of variables. Conceptually this idea is rather obvious; however, it is generally quite difficult.

Two distinct categories of data that are important in the context of development planning and decision-making are natural resources data and socio-economic data. Often, these two categories of data have been regarded as separate entities in so far as collection, handling and presentation are concerned. One reason for this is related to the differences in the data as such. For example, data on natural resources lends itself better in general to spatial characterization than socio-economic data. Another reason is the traditional sectoral approach to planning and decision-making.

It is evident, however, that integrated approaches to planning and decision-making provide a strong rationale for the integration of natural resources data and socio-economic data.

In the design of information systems that are intended to integrate different sets of data, each having different concepts and formats, it is recommended that one distinguishes between 'basic' and 'related' sets of data. As will be discussed later (Chapter 4), a related data set may be linked to a basic data set by applying the concepts and formats of the basic data set.

For the integration of natural resources data with socio-economic data, data sets on natural resources will often be selected as basic spatial data sets. Aerial photography and other remote sensing techniques, which record natural resources data, may be applied in the design and operation of the envisaged information system. This will be discussed later in greater detail (see Section 4.1).

2. GENERAL CONSIDERATIONS RELEVANT TO INFORMATION SYSTEMS

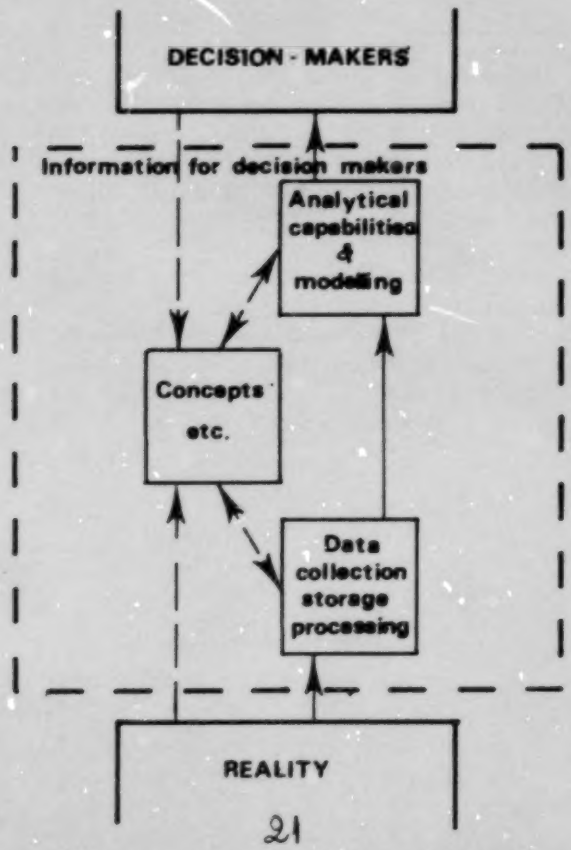
Geographic information or (geo-information) systems are special cases of information systems and should therefore take account of proven practices in information system design and operation. Since numerous studies have dealt with this subject only the most relevant themes (relative to geo-information systems) will be briefly treated here.

In setting up an information system, a broad view should be taken of the system which includes as components a conceptual base (i.e. the definition of data items and of the relationships between them), procedures for data collection, storage and processing as well as capabilities for analysis and modelling (3). Data collection, storage and processing require a properly defined conceptual base since an inconsistent definition of data elements may result in data which is of little use. Analytical capability and modelling in turn closely depend on the other two components.

BEST COPY AVAIL

GUIDELINE 1: A balanced development of the information system and of its components (conceptual base; data collection, storage and processing; as well as analytical capabilities) should be aimed at.

A schematic representation of the components of an information system is given in Fig. 3 below.



21

Fig. 3: Components of an information system (after Bonnen, 1975).

2.1 IDENTIFICATION OF THE USERS AND USES OF INTEGRATED SPATIAL INFORMATION

The central issue of the present report is how to integrate spatial information. Before this issue can be treated, some preliminary questions have to be considered dealing with the context of the envisaged information system. For instance, does the envisaged information system have to support national, regional or urban planning and decision-making functions? Does it have to deal with a single- or a multi-sectoral environment? The local context is particularly important in cases where methods developed in advanced and industrialized countries are considered for implementation at a new site - perhaps in a developing country.

Some typical situations where spatial information is needed are:
- conservation and exploitation of natural resources;
- development of land use plans;
- urban and regional development planning;
- monitoring of changes in land cover (i.e. changes in forest cover, changes in crop management practices);
- monitoring of environmental changes (i.e. deterioration of soils, pollution, hazards like floods);
- siting problems for specific enterprises or land use practices;
- assessment of food/nutrition production.

The above examples demonstrate clearly that most information required for comprehensive planning and decision-making includes both natural resources data and socio-economic data which must be presented to the user in a compatible manner. From these examples, it will also be understood that the required information is usually only available from multiple sources. The underlying (raw) data usually reside in several departments, often in different formats and scales and thus require considerable planning, time and effort to be integrated.

2.1.1 ORIENTATION OF AN INFORMATION SYSTEMS

The orientation of an envisaged information system has to be understood in advance. Two principal system orientations can be distinguished, depending on the kind of 'mission' the information system is expected to fulfil (4):

- general purpose or supply-oriented systems are systems of which the primary purpose is to supply data and/or information which can be used by a heterogeneous group of users. Consequently, these systems are not specifically oriented towards the activities carried out by the users. The information services of the central statistical bureaux are classic examples in many countries.

- specialized or function-oriented systems are those which serve and are geared towards a rather well-defined group of users.

Although ministries are generally function oriented, they may host both function- and supply-oriented information systems. A good example is the Ministry of Agriculture and Forestry in Uganda. In this Ministry, a purely supply-oriented system has been set-up which services the Ministry's Planning Unit; the various Government departments as a whole plus international agencies, such as World Bank, International Monetary Fund (IMF) and Food and Agriculture Organization (FAO); and, more recently, the various donor countries. Each of these requires information on the agriculture and forestry sectors, but their requirements differ in composition, detail and level of analysis. The system has therefore been made sufficiently flexible to cater to all these needs.

On the other hand, the information systems set up for the Research Stations of the same Ministry are function-oriented in that they provide information needed to determine combinations of the various inputs and conditions which are best for a given crop or are otherwise geared to fulfil specialized needs.

Since their end-uses are very heterogeneous, supply-oriented information systems require much more flexibility than function oriented systems, and this need for flexibility increases the difficulty in deciding on the subject matter of the system as well as the number and the formats of the various sets of information to be produced. The data from supply-oriented information systems often requires some further processing and analysis in order to be of use to the end users. It may be useful to envisage a situation where function-oriented systems supplement and make use of supply-oriented systems, the latter being designed for overall national, regional or local needs.

2.1.2 IDENTIFICATION OF THE USERS

In general, two different approaches can be followed in identifying users of information:

- an inventory be made of the existing flows of data and information, as well as the utilization of data and information;
- an inventory be made of relevant decision processes and the potential users of information.

Either of these inventories can be confined to the specific sector of the information system in question.

By analysing the existing flows between producers and users of data/information, a more or less complete list of present users may be obtained. It should be noted that such an approach describes the existing situation. The identified flows correspond to information requirements within the existing administrative structures.

The evaluation of the existing situation should be supplemented by a study of the needs for information including both expressed and latent needs. Through study of the relevant decision processes, the users participating in these processes and needing information are identified.

This approach concentrates on the context of decision-making and opens up the possibility of evaluating (in collaboration with the users of information) data flow structures which have previously been identified.

It is clear that through these two approaches two groups of user communities, in general not identical, will be identified. By applying the two approaches together, a more complete coverage of the existing and potential user community may be obtained, and it should also be possible to get an idea of the degree to which the present data needs are satisfied.

It should be understood that the users may be involved at different levels in the decision-making process which should be considered separately in analysing information requirements. For example, the following three levels can be considered in the planning process (5):

- policy planning, which should keep long-term considerations in view, and which is primarily concerned with opening up new questions;
- strategic planning, which generally has as its main function the presentation of choices;
- operational (executive) planning, which is short-term decision-making concerned with developing a plan of action.

A problem in identifying users of information, particularly in developing countries, is the limited awareness of the potential role of information. This means that many of the potential users and many demands may not be easily identified. It may be necessary, therefore, to apply an iterative approach; i.e. to regard initial lists of users as a first approximation and gradually arrive at a comprehensive list.

The approach for the identification of the user community will differ depending on the orientation of the information system:

a) for general-purpose, supply-oriented information systems, which are not geared towards the execution of specific functions, emphasis should placed on the inventory of the existing flows of data and information, to be complemented by studies of the relevant decision processes.
b) when dealing with function-oriented information systems, the emphasis should be on the analysis of the particular decision processes, to be complemented by an inventory of flows of data.

For several reasons, including overall system efficiency and relevance, motivation of user interest and avoidance of suspicion, users should be involved in the design and implementation of an information system. The individual users are often difficult to identify in advance, particularly in the case of supply-oriented information systems, and representative bodies of the user community may then be formed to advise on content and design of the information system. The execution of a well-chosen pilot survey may serve not only to identify users of information but also in making them aware of the potential impact of information systems on their professional activities.

GUIDELINE 2: In order to identify the (potential) users of information, it is suggested that inventories be made of:

- existing flows and utilization of data/information;
- the relevant decision processes and potential information users at different levels of responsibility.

These inventories can be confined to the specific sectors and/or user communities of the information system in question.

The emphasis in approach will depend on the orientation of the information system, i.e., general purpose, supply-oriented or function-oriented.

The execution of a pilot survey in collaboration with the user community may facilitate a subsequent full inventory of the needs of users and meaningful user participation in the envisaged system.

2.1.3 UTILIZATION OF INFORMATION

Before commencing the design of an information system the situations in which the information is to be utilized should be clearly understood. In a wide variety of sectors, including rural and agriculture development, physical planning, environmental management, resource development, engineering, etc, the following basic cases are encountered:

1. research situations (formulation of theories or testing of hypotheses);
2. inventories and other studies as a basis for new activities, policy orientations, projects, and programmes;
3. monitoring and/or evaluation of the impact of certain activities.

In any of the three foregoing situations the "informal" information component (including ideas, 'hunches', feelings and experiences that are circulating amongst the people involved - not only decision-makers, planners, extension workers and the like, but also farmers, entrepreneurs, etc.) can be very important. In such cases formal information systems may still be useful in confirming and providing quantitative back-up ('underpinning') for the existing informal information.

The above types of information use should be regarded as "idealized situations"; in practice, most information systems deal with a mixture of them. The requirements placed on the utilization of information for research, inventories, monitoring/evaluation and 'underpinning of informal information' differ considerably as will be pointed out in the following sections. In each case it is extremely important to determine the minimum level of accuracy, precision and resolution required in each data set.

GUIDELINE 3: Before considering the design of an information system, the types of utilization situations (e.g. research, inventories, monitoring/evaluation, 'underpinning of informal information') and the required degree of accuracy, precision and resolution in data must be identified.

2.1.3.1 RESEARCH

The utilization of information in the formulation of theories is a scientific exercise. Its ultimate goal is to gain knowledge or insight into a particular relation between effects and causes (6). The development of theories requires a broad range of variables but the number of cases can usually be limited. For example, the development of urban and regional theory may involve a programme of special survey work and systematic research in a limited number of carefully selected cities and regions (7). When developing or using a model in research, the particular model determines the data needs, and thus the design of the content of the information system. The information base used in the formulation of theories (e.g. models) may thus differ considerably from the information base used in subsequent practical applications of the research.

2.1.3.2 INVENTORIES AND DESIGN STUDIES

Information for the design of new activities, policy orientations, projects and programmes should provide a consistent description of reality. This information is usually included in general inventories (base line studies) in which coverage and timeliness rather than detail are emphasized, and in more detailed design studies for particular projects.

Information systems for these situations may serve a heterogeneous group of clients. Therefore the structure of the information system should be designed in such a way that it permits storage and retrieval of data at various levels of detail. Such flexibility can be greatly facilitated through an appropriate standardization of concepts and formats at data collection stage.

GUIDELINE 4: Information systems serving a heterogeneous group of clients/users who face different situations should be flexible and adaptable to permit easy inclusion and retrieval of detailed information when needed. This generally requires some standardization of concepts and formats for data collection.

2.1.3.3 MONITORING AND EVALUATION

When implementing activities such as projects and programmes, the need arises to analyze their progress and assess their effects and impacts (8). Two sub-categories of such assessment can be recognized:

a. - the follow-up of a process (e.g. project implementation) with respect to the resultant changes and effects, which is referred to as monitoring, and
b. - the assessment of project/programme achievements on the basis of anticipated objectives/targets, which is referred to as evaluation.

A supporting information system for monitoring has to provide timely information on a limited number of key indicators. Often, it can be based on simple approaches to collecting and handling a minimal amount of information. Its content is designed to assist policy makers in choosing options for corrective action.

The term "monitoring" is often restricted to (specific) projects and programmes. However, monitoring can be applied in an extended sense as well. For example, a government agency may wish to follow one or more geo-processes such as soil erosion, land degradation, deforestation, changes in wildlife stock, crop growth, flooding or effects of urbanization on the environment. The monitoring of geo-processes deals, by its very nature, with spatial phenomena. Therefore, the supporting information systems have to treat the spatial dimension adequately.

Methodological aspects of monitoring, such as reliance on base-line studies and/or control groups, should be considered carefully before the corresponding information system is designed. Insight is required into the implementation process as well as into the related decision-making process before the data to be collected can be identified. Therefore, it is recommended that the design of any project or programme include the design of the supporting monitoring and information system as well.

The design of an information system for monitoring by a specialized government agency, should be embedded in the management and decision-making structure of that agency. In the case of an agency which does not take remedial actions itself, the design problems of information systems for monitoring are similar to those of general supply-oriented information systems.

GUIDELINE 5: Information systems for monitoring or evaluation of specific projects and programmes should be embedded in the design of these projects and programmes, whereas more general information systems for monitoring or evaluation should be embedded in the management and decision-making structures of the information users.

The monitoring of geo-processes requires a thorough understanding of these processes as well as of their favourable and adverse effects This implies that criteria for levels of magnitude, frequency and incidence over space must be known and that key indicators must be established to measure them.

GUIDELINE 6: Monitoring requires the establishment of sets of key-indicators and an information system that can adequately cope with them.

2.1.3.4 UNDERPINNING OF 'INFORMAL INFORMATION'

Most day-to-day decision-making and planning, particularly at the sub-national level, relies on 'informal information'. A general problem with 'informal information' is that it is often qualitative, subjective and difficult to collect and verify. However, 'informal information' may well prove to be more accurate than offical estimates in a given domain, for example in the case of crop-yield data provided by local farmers.

The relation between 'informal information' and formal information systems is two-fold:

a. 'informal information' can supplement sets of data, thus leading to more complete information. In this sense, it is considered as a source of data;
b. 'informal information' may in itself give a good idea of the relevant issues but may have to be checked against precise data before embarking on appropriate action. For example, an experienced planner may 'know' about a certain change in agricultural practice but he needs more quantitative and objective information. Here, the formal information system has to verify and improve upon the already available 'informal information'. This can be called underpinning or rationalization of 'informal information'.

GUIDELINE 7: When designing an information system, the availability of 'informal information' should not be overlooked; indeed, such information should be utilized to the maximum extent possible.

2.2 DESIGN CONSIDERATIONS OF INFORMATION SYSTEMS

2.2.1 EXISTING INFORMATION CAPABILITIES AND PRACTICES

Before designing an information system, it is advisable to take stock of existing data and information, information capabilities and practices. With regard to existing data and information, the important considerations include the kinds of data (e.g. topographic and cadastral, geological, population and social service distribution), the storage of data (e.g. in office files, published and unpublished research reports, annual and special reports), and the formats and scales in use (e.g. standard map projections and scales, administrative area maps, census tabulations and remote sensing maps and data).

This survey may be presented as an annotated list of data sources at the national level. As part of this exercise, a preliminary list of data gathering agencies should also be compiled.

Data produced through non-indigenous consultancy arrangements should be included in the survey as well. Such data are often collected for inventory and design purposes, with natural resources data (e.g. soils, water resources, wild life) frequently being the major component.

It is also important to ascertain the indigenous information-handling capability in terms of equipment and manpower. This will be useful in planning, programming and budgeting for the proposed information system, particularly in the areas of personnel training and equipment allocation.

Finally it is very useful to identify existing information management practices relating the custody of particular sets of data/information and their accessibility. It sometimes proves difficult to obtain specific data either because they are located in large organizations dealing with a variety of data sets or because the gathering agencies overlap in their spheres of activity. Legal constraints arise in a number of situations (e.g. in the case of classified information). Often, however, the problems of accessibility relate to bureaucratic control. For example, even in cases where data collection is the duty of a centrally mandated agency or its affiliates (e.g. topographic mapping and census data) the statutory provisions of the agency usually permit accessibility to data from such an agency.

The above survey of data sources, capabilities and practices could also be very useful in identifying the relevant elements of a national, regional or local data and information plan (9). In most developing countries, such a plan does not exist at the moment. The establishment of a government data and information plan could form the starting point in any programme for information system development.

GUIDELINE 8: The design and creation of an information system should be placed in the context of other information systems, existing as well as anticipated.
To this end, a survey of existing data and of information-handling capabilities and practices should be made before embarking on the design of a new system.

2.2.2 HIERARCHICAL SYSTEMS AND DECENTRALIZED INFORMATION SYSTEMS

It often proves to be advantageous to design an information system in a hierarchical way with sub-systems for specific aspects or levels. Such sub-systems could be designed for various sectors, such as agriculture, industry and transport, or for different geographic levels, such as the local, regional and national levels. These sub-systems are generally inter-linked to permit data aggregation and/or cross-referencing of data.

Whenever possible, the hierarchical structure of the information system should reflect existing hierarchical structures in the processes of policy/decision-making in order to facilitate the organization and flow of information.

The integration of the various (hierarchical) sub-systems can be done manually or with the aid of computers The functions in the linkage procedures must be made explicit. Basically two alternatives can be considered:

- decentralized data collection at a number of sub-national levels versus centralized collection with sole responsibility at the highest level;
- decentralized processing of information at a number of sub-national levels versus centralized processing at the highest level only.

An example of a highly decentralized information system can be found in Uganda where the Ministry of Agriculture and Forestry has District Statistics Offices in each of the districts. Each of these units collects and partially analyzes the information on district-level agricultural activities for both local and headquarters use. At the same time, the district unit supervises the various activities of the Annual Sample Survey -- an exercise designed to provide more reliable data on the agricultural sector. Both sets of information are passed on to the headquarters of the Ministry for further processing, storage and dissemination. The analyzed information is finally passed on to the Central Statistics Office.

Due to the fact that the District Statistics Office supervises a relatively small area and that local personnel are used, it has been found that farmers and the various sub-district offices co-operate easily in the collection of data. This might not be the case if the Central Statistics Office personnel and control were to be used.

It is also possible to envisage a partly decentralized system in which the responsibility for either data collection or data processing is centralized. In the former case, extracts of the central data base could be provided to lower level agencies for use in re-aggregation and information services whereas the latter case information could be collected at a variety of levels but would be processed at a central facility.

A major difficulty is that, unless processing is done very quickly, the time lag between data collection and information supply may be so long that the information may be out of date before it can be used. This may necessitate transferring preliminary processing capabilities to lower levels in order to speed up the delivery of information. If preliminary aggregation is carried out at the lower levels, local decision makers can use these approximations as required. At a later stage, the estimates can be revised at a higher level and the authoritative information made available at all levels of Government.

Decentralized systems may also facilitate public participation by facilitating accessibility to information. A prerequisite for this is that the public be involved in the process of planning at the various levels and sectors. Experiences in many countries show that the local/district level is the focal point for participation of ordinary citizens in development.

Decentralization of data collection and/or processing can improve the relevance, quality and accessibility of information for local decision-making. This is particularly important within a context of multi-level planning and decentralized decision-making For example, the need for establishing decentralized information systems arises from the adoption by many developing countries (as India, Sri Lanka, Kenya) of concepts of "integrated rural development". Modern technology, and especially small computer systems, may be particularly appropriate for decentralized information systems.

Decentralization may prove more costly than the centralized approach and in this case, the justification for it must be sought in increased effectiveness or benefits of the decentralized system. In addition to the several advantages outlined above, it should be noted that the parallelism of decentralized systems may be seen as ensuring a certain robustness or capability to fall back on, which should be weighed against added costs associated with decentralized processing.

Distributed systems rely on effective co-ordination between co-operating units. Procedures must be developed to ensure the maintenance of data integrity and the support of multiple user groups.

GUIDELINE 9: Decentralized information systems in general have a great potential for facilitating local and sectoral decision-making. Inexpensive data processing technology is becoming available to implement such systems. Decentralization, however, requires careful formulation and regulation of the linkages between the central unit and the peripheral nodes as well as a strengthening of local capabilities.

2.2.3. SKILLS

Training programmes should not be restricted to the short period of setting up an information system, as is too often the case particularly when donors are providing advanced technology. Also needed are longer-term programmes for various information specialists and users at different levels. The considerable resources required for the establishment and operation of such training programmes are often overlooked.

Processing of data may be particularly difficult at the local level in many developing countries, due to a lack of trained and qualified personnel. In such cases concerted efforts may be necessary to strengthen the planning machinery at the sub-national level, including manpower and facilities for data processing, as, for example, is currently being attempted in India.

GUIDELINE 10: At an early stage in the design of an information system, training programmes have to be developed to guarantee the availability of skills. The resources needed for these training programmes must be ensured

2.2.4 TECHNOLOGY

The choice of the appropriate technology for an information system should be considered as an integral part of the system design. Even if a great deal of information technology can be imported, it is usually prudent to centre information systems around the types of equipment and procedures that are available and known in the local setting. The tendency to continuously import the latest technologies may prove to be a waste of resources and place many systems at a risk if the importing capability of the country decreases.

The assessment of the appropriate level of technology required is very complicated. This is mainly due to the fact that technological possibilities are developing rapidly and almost continually. Therefore, not only should the local situation be analyzed ("what is wanted and what is locally available") but technological developments, both national and international, should be monitored In several areas technology has developed in such a way that it comes within reach of many developing countries without the necessity of following the same "learning process" as the industrialized countries. For example, inexpensive micro-computers are within reach of most developing countries these days, as is borne out by the success of the geodata handling system developed by HABITAT, which is especially suited to local/district use (10).

GUIDELINE 11: The envisaged information system should take into account both locally available technology (including hardware and software as well as manpower) and international trends in technological development.

2.2.5 CHANGE STRATEGY

Setting up an information system is not a one-time activity. Changing circumstances require a continuous process of re-design and adaption for which a change strategy should be developed (11). Lack of compatibility between successive computer systems (comprising both hardware and associated software) places a burden on data collection as well as adding to costs for hardware and maintenance. Quite often, once a system has been set up and put into operation, data collection procedures, become dependent on that system in order to make optimal use of all its potentialities. When a new computer system has to be selected, the choice is restricted in practice to those that accept data collected according to the previously fixed procedures since others would require major changes in data collection and make existing data more or less obsolete. When setting up an information system, it is prudent to develop data collection procedures, concepts and formats that are as independent as possible of the pecularities of particular system hardware and software adopted. The advantages are that, at the expense of not fully utilizing the present system's capabilities, the system may change rather smoothly in the future without affecting data collection too much.

GUIDELINE 12: In parallel with the selection of a computer system, a careful strategy for change and adaptation should be developed. Data collection procedures should depend as little as possible on the system's hardware in order to facilitate necessary future changes.

In designing an information system, stock should be taken of existing information handling practices of the users in order to:

(1) determine the types of data-handling activities that are adequately performed by manual methods;
(2) identify activities that are too complex to be readily handled by computer methods, i.e. those which are adequately handled by intuitive mental models;
(3) identify the middle segment of data-handling activities that are amenable to improved manual methods or computer-assisted methods.

The assessment of existing practices should also include data element definition, formats, etc., which sometimes need revision or standardization. Carefully formulated procedures can help greatly in ensuring the effective use of technological capabilities.

GUIDELINE 13: It is generally preferable that information systems develop within existing administrative and institutional contexts. It is usually advisable to examine the evolution of manual methods into computer-assisted methods. These improvements should show a benefit as soon as possible in order to sustain the confidence and interest of the users.

3. COLLECTION AND HANDLING OF SPATIAL DATA; GEOCODING

3.1 ALTERNATIVE METHODS OF GEOCODING

Geographic referencing (or geocoding) is the assignment of a spatial reference/identifier to spatial phenomena (often, but not necessarily, restricted to topographic features). As will be shown in the following paragraphs, alternative methods of geocoding can be applied (12).

The need for geographic referencing will be more strongly felt when emphasis is placed on physical planning at regional and local levels, where data relating to different sectors and originating from different sources have to be integrated.

The choice of the appropriate method of geocoding depends on the requirements for information with a spatial dimension. It is important to express as clearly as possible the (analytical) purpose for which the spatial data are required. The following types of analysis should be considered when formulating requirements for spatial information:

- search for attributes of a specific location (area, line, segment or point)
  - e.g. - land use for a particular plot
    - household composition for a particular dwelling
    - capacity of a transportation link

- aggregation of attributes over a defined area
  - e.g. - total number of inhabitants of an administrative area
    - total amount of fertilizer actually applied within a particular area
    - total acreage under cultivation in a particular area

- surface calculation; delineation of homogeneous areas (areas with similar characteristics)
  - e.g. - The boundaries of administrative units each of which is required to have a normalized spatial distribution in terms of density value, e.g. population per sq.km.

- distribution analysis; distribution of certain phenomena over time and space
  - e.g. - deterioration of soils due to erosion
    - deforestation
    - suburbanization

- logical operations; identification of areas or points which satisfy a logical condition ("if this and/or that, etc.")
  - e.g. - areas with a soil depth over x meters and a slope less than y
    - areas with high population density and low literacy rate

- location analysis; determination of the location (of a point or a line) as a function of certain phenomena with spatial dimensions
  e.g. - location of schools with respect to student population
  - catchment areas of facilities
  - optimal location of a road

- flow analysis; through networks or between regions
  e.g. - traffic along a pattern of roads
  - inter-regional migration

- diffusion models;
  e.g. - spread of innovation
  - spread of disease and pollution

As a supplement to the above specific calculations it may also be necessary to envisage:

- the application of general analytical methods
  e.g. - spatial autocorrelation
  - spatial spectral density analysis.

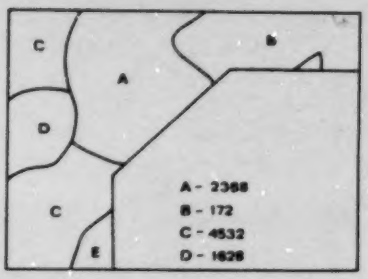
The tools to meet the users' requirements rely heavily on geocoding. Objects or phenomena can be given a spatial reference directly (direct geocoding) or they can be related to other objects with a spatial reference (indirect geocoding). An example of indirect geocoding is an individual person coded by the spatial reference of the building in which he lives. Direct geographic referencing is usually carried out by one or more of the following methods which are shown in Fig. 4.

- definition of areas. The total area concerned is sub-divided into smaller areas (units). Spatial information is referenced by these smaller units. Basically, two different modes can be applied in this respect:
  (a) - delineation of (existing natural or physical) topographic areas or
  (b) - creation of arbitrary terrestrial areas through a network of grid-cells.
  Topographic areas can be delineated by defining their respective boundaries. Usually, these boundaries are defined as polygons with points of known co-ordinates. Examples of topographic areas are administrative areas, enumeration districts, parcels, surface areas of buildings and geographical land units;

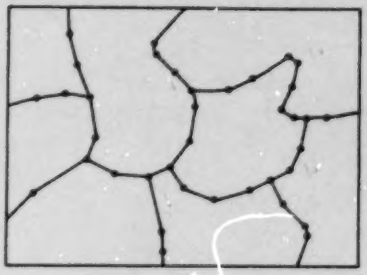
- definition of lines, representing topographic line-features. Examples are roads, pipelines, etc. Usually, these line-structures are defined as polygons with points of known co-ordinates;
  segments are special cases of lines. A segment is defined as a length of a line between two distinct nodes, thus enabling the representation of a street network for example. In addition to representing lineal features (roads, sides of areas, etc.), information about the topological relationship between the given line and the the areas on either side of the line is also captured to greatly facilitate automated editing as well as modelling of network processes.

The spatial information in the system (e.g. traffic density) may refer in such cases to a topographic line-feature or part of it;

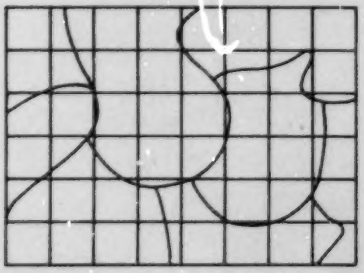
- definition of points. Point co-ordinates may represent specific topographic areas in the form of centroids of parcels, buildings, etc., or representative locations for spatial phenomena like boreholes, observation sites, checkpoints, etc.



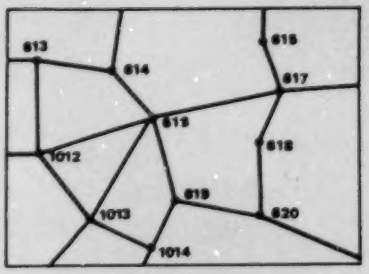
Delineation of topographic areas based on their real boundaries.



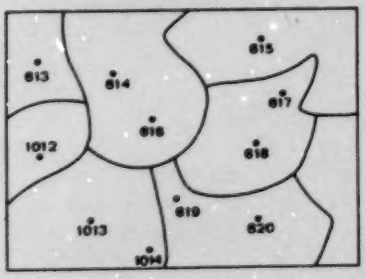
Delineation of topographic areas by defining their boundaries as polygons.



Creation of arbitrary terrestrial areas by a network of grid-cells.



Definition of line-segments.



Definition of points at (x,y) co-ordinates.

Fig. 4: Different methods of geographic referencing (after Tomlinson, 1972).

If the purpose of the geographic information system is mainly to keep administrative records, such as population or property files, the objects may receive a direct geographic reference (e.g. by representing parcel boundaries or centroids by co-ordinates) or indirectly (e.g. by relating them to postal areas or street blocks).

The geographic references, whether direct (viz. areas, lines, segments, points) or indirect (viz codes, names, addresses), are referred to as basic spatial units.

At present, a number of major categories of software for handling spatial information are available:

- engineering mapping systems for photogrammetry, geodetic data, topographic basemaps, road engineering maps, utility maps, facility management, etc.;

- geographic base file systems for management of information on streets or similar networks and the space they enclose (urban management, statistical mapping, census area mapping, etc.);

- property or parcel information systems (for taxation, land cadastre, etc.);

- generalized thematic and statistical mapping systems (for natural resource inventories and management, regional planning, census mapping, etc.);

- data systems which catalogue administrative records or bibliographic data sets related to geographic objects (population registers, address gazetteers, property files, etc.);

- image processing systems (typically associated with the processing of satellite image data).

3.2 APPLICATION OF DIFFERENT METHODS OF GEOCODING

There are several reasons why different methods of geographic referencing should be considered. One important consideration is the volume of data to be handled and another lies in the spatial characteristics of the data, i.e. the degree of spatial variation. In this respect, it is meaningful to distinguish between two classes of data base in choosing a geocoding method (13):

- the regional inventory;
- the local inventory.

Regional inventories typically do not deal with locations of individual parcels or other features like utilities and facilities. The spatial variation of the phenomena considered can be dealt with in an aggregated and generalized manner. For example, it is not the crop production per individual parcel that is of interest but, instead, the total or average production for a larger area. This 'larger area' forms the basic spatial unit for which relevant data have to be collected and spatial homogeneity is assumed within this basic spatial unit.

The following major types of basic spatial units can be recognized for regional inventories:

- administrative areas, such as village and planning districts. These administrative areas normally have well-defined boundaries. An advantage in using this type of unit is that many administrative records are kept and statistics collected at this level;

- natural or ecological units, such as ecological zones and watersheds. Most of these units do not have well-defined boundaries. For generalized natural resource inventories, natural or ecological units may form a basis for stratified sampling. A first approximation in delineating such units may be obtained from satellite or other remote sensing techniques;

- system of grid-cells. The area is sub-divided into a system of grid-cells, each with well-defined boundaries. Relevant spatial data or information refer to a particular grid-cell. This type of basic spatial unit may apply to generalized rural or regional inventories as well as urban inventories. In rural/regional inventories, the Universal Transverse Mercator (UTM) Grid may be suitable.

Local inventories often deal with individual objects or elements. As such, the relative location of these objects is of importance. Compared with regional inventories, local inventories generally deal with a higher level of spatial detail and, consequently, with greater amounts of data per unit area.

The representation of spatial units for local inventories can be done by encoding the boundaries through:

- topographic mapping, representing the boundaries of relevant topographic features or
- cadastral mapping, representing boundaries of different land properties.

For both of these types, the vertices of the boundaries are given in a known system of co-ordinates. Alternatives to these types of geo-referencing include:

- representation of local inventories by centroids given in a known system of geographic co-ordinates;
- geographic base files, which make use of a geographic frame associated with a suitable topological structure, e.g. a street network

In geographic base files, the spatial units mapped are street blocks, bounded by street segments. (The latter being the 'basic spatial units'.) Each of these street blocks receives a unique code and can be used for a variety of administrative and management purposes. The US Bureau of the Census developed geographic base files along these lines to geocode decennial censuses starting in 1970. These geographic base files may be used for automated address coding, spatial analysis and computer mapping of socio-economic and demographic data; data can be assigned to street blocks or other spatial units such as traffic zones or planning districts. Both the US and Canadian Bureaux of the Census are developing techniques to derive enumeration maps from such geographic base files. Other census bureaux and local authorities are applying geographic base files as well. In several countries, for

example, it is common to have cadastral files on which each property is given a hierarchical code consisting of the administrative zone code, street block code and property code.

Geographic base files also provide basic spatial units to which more detailed data can be aggregated. For example, data on individual addresses can be aggregated to the street block to which they belong. These aggregated data can subsequently be displayed and/or analysed.

Although a particular method of geographic referencing can often be used simultaneously at several different levels of spatial detail (for example the referencing of data sets to areas corresponding to different administrative levels such as the province, the municipality or the quarter to which the data pertain), it may well turn out that different methods of geographic referencing are required for the various data sets; for example, some data may be collected at the level of municipalities whereas other data are collected at the level of land units (parcels). Conflicts may appear with respect to the preferred geocoding system for integration of these data sets. For example, conflict might arise over the decision to select a grid-based over a polygon-based geocoding system. It is usually wise not to choose a single alternative but to design the information system in a modular way so that different sub-systems may be developed according to the various geocoding systems needed in the particular context.

From the view point of managing such complex, modular information systems, it is preferable to keep the number of different geocoding systems to a minimum and standardized concepts, formats and geographic referencing should thus be implemented to the extent possible. This implies that a few major users of geocoding systems may have to adopt and apply data collection methods which are not optimal for their specific application.

If the various data sets exist in the form of ordinary maps (or map-like products) based on homogeneous mapping units with identified boundaries (e.g. soil maps), the integration of these data sets follows the same logic as superimposing transparent overlays on top of each other. Two alternatives to manual overlay of real transparencies are:

- the definition of a single-grid system (common to all the map-like data sets) and the subsequent determination of the content of the grid-cells for each data set, and
- the use of digital polygon overlay techniques.

Modern equipment facilitates the automated polygon-to-raster conversion (often with very fine grid-cells) and the overlay of polygon-based data sets represented as co-ordinates. If these transformations have to be carried out by manual methods, the resulting grid-system will be much coarser than the result that could be obtained with automated approaches.

3.3 CHOICE OF THE APPROPRIATE SYSTEM OF GEOCODING; TRADE-OFFS

The appropriate method of geographic referencing depends on the particularities of the individual system concerned. However, some general guidance can be given in terms of three major considerations:

- spatial variability of the phenomena concerned;
- possibilities for data collection;
- the decision-making context.

3.3.1 SPATIAL VARIABILITY

The size of the optimal basic spatial unit tends to decrease if the spatial variability of the phenomena considered increases. However, in some cases a particular utilization of the data may be satisfied by providing a measure of the spatial variability within a particular area instead of the actual spatial distribution, thus allowing for relatively larger basic spatial units.

If, for a given application, one has rather large basic spatial units, sampling procedures may be considered to obtain the data that are to be assigned to these basic spatial units. In East Africa for example, the Systematic Reconnaissance Flight (SRF) is commonly used. This is a sample survey technique for quantifying the spatial distribution of land-use parameters through a combination of visual observation and vertical photography from low-flying, high-wing light aircraft (14). One or more vertical sample photographs are taken within each (planning) unit and the data extracted from these (sample) photographs are referenced to the (planning) units concerned. The Universal Transverse Mercator (UTM) Grid, which is easily projected onto the aerial photographs or other cartographic/thematic data sources, is normally used for such sampling.

GUIDELINE 14: The optimal size of the basic spatial units of an individual data set depends on the required spatial precision of the phenomena considered. Relatively large basic spatial units are often acceptable, thus simplifying the sampling procedures needed to collect the appropriate data.

As a whole, the optimal level of spatial detail for a system, or sub-system, is a trade-off with respect to the requirements for each of the individual data sets. For example, if a network of grid-cells is adopted for geocoding, the optimal size of the grid-cells should be such that administrative units as well as natural land units could be represented effectively by compositions of grid-cells. Often, the optimal level is a smaller common spatial unit than needed by any of the individual data sets. Considerations in this respect include detail required and efficiency in data handling. In situations where high levels of detail are demanded, one should critically examine whether such detail is required for all of the data sets. Often this is not the case. Data sets with exceptionally high levels of detail could be regarded as separate but co-ordinated data sets.

GUIDELINE 15: The optimal level of spatial detail for a geographic information system as a whole represents a compromise with respect to the requirements for each of the individual data sets. As a means of improving processing and storage efficiency, individual data sets with exceptional degrees of detail can be regarded as separate but co-ordinated data sets.

With modern equipment, it is now technically possible to devise systems which can handle data at varying levels of detail. However, this does not mean that stored data can be integrated without limitation. The final level of spatial detail of the information to be provided will depend on that of the different data sets used. This implies that, even if the level of spatial detail of only one data set is below the required level, the required final level of detail can not be obtained. On the other hand, it will in general be possible to work with a lower level of spatial detail than that employed in the data sets by aggregating the data as appropriate.

In general, it is advantageous to carry out geographic referencing in the greatest possible detail in order to ensure the greatest potential for subsequent application of the data. However, as geocoding becomes more detailed the collection and processing of the data will generally be more costly in terms of finance, technology, skills and organization.

In smaller or industrialized countries it may be preferable to use topographic features in order to locate spatial phenomena, whereas in large and predominantly rural countries small administrative units (districts, municipalities, villages, parishes, etc.) may form the most convenient method of geo-referencing.

3.3.2 POSSIBILITIES AND LIMITATIONS IN DATA COLLECTION

Limitations in data collection often do not allow high levels of spatial detail. For example, collecting environmental data in terms of individual parcels is a costly and time-consuming process. In general, the maintenance of high levels of spatial detail in data collection involves high costs which have to be weighed against the benefits.

A rather related question is the degree to which current methods in data collection should be considered as fixed procedures. In other words, one must consider whether it is possible to adapt data collection to a given method of geographic referencing or whether the method of geographic referencing has to be selected to conform to an existing method of data collection.

3.3.3 CONTEXT OF DECISION-MAKING

Decision-making applications generally require a geographic information system based on either:

- a regional inventory (e.g. for regional natural resource planning)
  or
- a local inventory (e.g. for cadastre).

As explained previously (see Section 3.2), regional inventory information gives a global picture of an area, in terms of parameters such as population, total amount of crops grown, potential to increase acreage of production. Local inventory information, on the other hand, is often required by a city-planning or urban management agency which usually needs data concerning parcel-specific land use as part of its analysis.

Regional and local inventories differ considerably in the required level of spatial detail and require different kinds of mapping display (13). Most resource information systems are proposed and justified (and therefore funded) for generalized inventory purposes. As such, they are valuable planning tools, but their internal logic and coarse method of geo-referencing (e.g. cell size) render them prone to serious misclassification of individual parcels of land. Regional (generalized) inventory information should not be mapped in a way that allows identification of individual land parcels. Nor should this type of information be made available for applications such as tax assessment. If the envisaged geographic information system deals with both type of data, viz. regional and local inventory, different methods of geographic referencing for the various data sets have to be considered.

In designing a geographic information system, the smallest spatial unit for which information should be made available is generally the smallest spatial unit in the decision-making process concerned. In case of regional- or district-level planning, this is often the village. For analytical purposes it may be desirable to use smaller units than the village, but in most cases these smaller units are only used as sample units.

In case of natural resources oriented decision-making, the smallest spatial unit for which information should be made available, will often be an ecological unit.

It is common to select a particular method of geographic referencing for comprehensive use in referencing all basic spatial units within the study area considered. For example, if the parcel has been selected as the basic spatial unit, all parcels within the study area are covered by this method of geo-referencing. For special applications, e.g. statistical analysis, a smaller spatial unit may be required as a sample unit. For example, within a parcel a few boreholes can be selected.

GUIDELINE 16: An appropriate method of geographic referencing can only be selected after determining the requirements in terms of treatment of spatial data, the required level of spatial detail and the need to integrate specific sets of spatial information.

4. ESTABLISHING AN INTEGRATED GEOGRAPHIC INFORMATION SYSTEM

4.1 INTEGRATION OF DIFFERENT SETS OF SPATIAL DATA

Integration of spatial information of different themes requires compatible geocoding e.g. for employment rates it is necessary that data on population and employment both refer to the same spatial units. This places specific requirements on the collection and presentation of data.

Basically, two different approaches can be followed when integrating spatial data.

One approach is to apply identical geocoding for the various data sets in the data collection stage. Conceptually this approach is simple and attractive, but it requires a large degree of technical and institutional/organizational sophistication in data collection and data storage procedures.

The other approach is to look for methods by which the geocoding scheme(s) defined for any given data set can be transformed (or adjusted) so as to correspond to the other schemes in use. This implies approximations and therefore a certain loss of information but takes into account that in practice most data sets are collected with different geocoding schemes.

GUIDELINE 17: Whenever practical, fully compatible geocoding schemes should be adopted at the stage of data collection Otherwise, methods have to be found to transform or adjust geocoding the geocoding schemes applied to individual data sets to permit matching with other data sets as necessary. This question should be considered with care since in specific situations mixed geocoding schemes can be more cost-effective than having a single scheme.

As has been mentioned earlier, a recommended strategy for the design of information systems intended to integrate different sets of spatial data, is to distinguish between 'basic' and 'related' sets of data. These two sets have an hierarchical relationship in that a related data set may be linked to a basic data set by the geographic referencing of the basic data set. This relation does not necessarily imply differences in the importance of the information concerned. Indeed, in terms of the priority of the decision makers, related information may be the most important.

In general, the basic data sets typically provide a more complete spatial coverage and a greater level of detail, whereas the related data set, which could be formed on the basis of sample surveys, will often have a much lower level of spatial detail or even only fragmented spatial coverage. Basic data sets provide a framework for extrapolation based on sampling.

The dynamic character of data sets is often described in terms of decay rate. In principle all data have a certain decay rate, ranging from low (e.g. data on bedrock geology), to high (e.g. data on political opinion). For the integration of spatial information from multiple data sources, data sets that have a high degree of variation over space and are not too dynamic in character are preferred as basic data sets. Examples of basic data sets are household files, plot files, or files on villages.

For the integration of natural resource data with different types of socio-economic and technological information, data sets on the natural resources will often be selected as appropriate basic spatial data sets.

GUIDELINE 18: An integrated geographic information system may be composed of separate but co-ordinated data sets each having a different method of geographic referencing. Of these data sets, one or more are selected as the basic data set(s) while the others are considered as related data sets. The basic data sets have a stronger influence on the methodology for integrating the various data sets.

Aerial photography and other remote sensing techniques can record natural resource data sets with relatively low decay rate (as in resource inventories) as well as data sets with a high decay rate (as in disaster monitoring). Obviously, the frequency of updating basic data sets should be proportional to the decay rate of these data if comparable accuracy is to be achieved.

Existing maps, photo-maps and mosaics, and remote sensing imagery can be used as basic data sets. Sometimes, interpolated map-like data sets may be regarded as basic data sets, e.g. soil maps, rainfall maps or population density maps; these set are often particularly useful in the design phase of an information system, providing that their scale and coverage are appropriate.

In Kenya, LANDSAT imagery was used in combination with ordinary aerial photography in land-use mapping (15). A preliminary map was produced by projecting LANDSAT colour composite transparencies onto 1:250,000 scale base map. The delineation of homogeneous land-use categories was based on the various colours, texture and patterns found on the imagery and subsequently checked from light aircraft and on the ground. The final mapping scale was at 1:1 million.

Sample aerial photography was taken within each major agricultural unit to give more detail and accuracy to the classification. Data from the analyzed photographs were then extrapolated to the whole unit and subsequently to similar units elswhere in the country.

This land-use map together with other environmental parameters such as climate, soils, human activities, etc., can be used in formulating a land-use policy. The map could also be used in urban and rural development, woody biomass assessment, rangeland planning and finally as a teaching aid in research and technical institutes. For the design of such systems, modern data collection technology in the field of natural resources should be carefully examined, especially aerial photography and other remote sensing techniques.

GUIDELINE 19: Certain kinds of data sets on natural resources which have a high degree of variation over space and a low decay rate are appropriate basic data sets within integrated information systems.

The quality of map data can be characterized in part by its level of generalization (in terms of scale and/or mapping intensity). For example, in the case of generalized natural resources maps, the location of the boundaries is often approximated. Consequently, the boundaries of areas resulting from integration (or overlaying) of various map-like data sets is also approximated. Another aspect is variability within individual mapping units. Although each mapping unit is typically assumed to be homogeneous over its entire surface, in general this is not the case.

The inherent standards, generalization and inaccuracies of the original maps and of the supporting survey procedures, should be taken into account when considering the integration of spatial data, since whenever data resulting from integration of generalized maps show a high level of detail, 'pseudo-facts' are likely to appear.

GUIDELINE 20: When integrating spatial data sets which differ in terms of scale and/or detail, care should be taken to inform the users of the inherent limitations of the data.

Integration of data sets with different geocoding schemes can be performed by a single designated data processing service. In general, the smallest spatial unit --whether this is a common spatial unit for the various data sets, or not -- may serve as an appropriate "vehicle" in this integration. Linking the different modules together can be done manually if no other solution is feasible. However, systems are on the market today which easily transform polygons into grids. A prerequisite is that the data-handling procedures (e.g. command languages in case of computerization) for each of the modules are mutually compatible.

GUIDELINE 21: The integrated geographic information system should be designed in a modular way. It may prove useful to foresee a common 'data transformation service' which may be either manual or automated.

4.2 CO-ORDINATION AND INSTITUTIONAL FRAMEWORK

To ensure effective development and use of integrated information systems, adequate linkages with the administrative structure are necessary. Experiences show that the best framework for this purpose is an Inter-Agency or Interdepartmental Steering Committee, composed of a variety of data gathering and user agencies, to plan, monitor and evaluate all aspects of the system's development and use. Although such a committee should bear some minimum degree of authority, its actual effectiveness depends on the degree of openess, division of responsibilities, objectives and interests among the parties involved.

The question of whether such a committee can function, given the proposed structures, should be carefully considered. It is often a wise policy to bring parties together on the basis of voluntary co-operation while maintaining existing identities; generally, this avoids too much suspicion and may stimulate a willingness to co-operate. When the envisaged information system deals with a single large domain (e.g. a sector or a geographical entity) the structure of the committee should be adjusted to the structure of committees responsible for other aspects within that domain.

The committee should ensure that required information is easily accessible and is not withheld because of "confidentiality" unless so classified under an appropriate legal procedure. Of course, the suppliers have to be protected against misuse of data and/or information as well. The use of information for a purpose it was not collected for easily leads to erroneous conclusions.

The establishment of an inter-agency committee may present several problems. Where a national information policy organization exists, the setting up of such a committee will be one of its major functions which should be carried out in consultation with the major users. In the case of many developing countries, such bodies may not exist, and another concerned agency may then be designated to convene and support the committee. Lack of co-ordination is often due to the fact that various existing survey and mapping organizations have developed independently over considerable periods of time, so that they have neither the possibility nor the incentive to adjust their production programmes to meet actual needs. In this situation the interagency steering committee could take the form of an 'Geographic Information Planning Board', housed within a general government planning agency (16).

GUIDELINE 22: An inter-agency or interdepartmental steering committee is an effective instrument to plan, monitor and evaluate the development and use of an integrated geographic information system. In such a steering committee both users and producers of spatial data should be represented. The committee could be housed within an appropriate government policy or planning agency.

5. POST-ESTABLISHMENT MANAGERIAL CONCERNS

Information systems need some form of control. Successful operation of an information system requires that its goals be specified and reformulated when required and that the various activities be co-ordinated towards the fulfillment of these goals.

When a geographic information system becomes operational, a number of post-establishment concerns come to the fore and must be dealt with to ensure continuous and successful operation. Several of these concerns have been introduced in various parts of earlier chapters. They are brought together here for review and re-emphasis as fundamental management responsibilities (17).

Properly planned and co-ordinated support structures go a long way in ensuring the success of the system. The absence of one or more support structures will easily erode the effect of those already set up and consequently result in the failure of the whole system. Leaving aside the essential element of political support (which was partially treated in Section 4.2), the necessary support structures can be divided into four areas:

- technical,
- financial,
- logistical, and
- managerial.

5.1 TECHNICAL SUPPORT

Support and maintenance of existing files, software and hardware are prerequisites. Particularly, backup files for data as well as for software have to be created to reduce the risk that information gets lost in the case of installation break-down. The creation of so-called 'shadow files' in different geographical locations is recommended. Whenever possible, backup of hardware should be secured through duplication of installations or by inter-agency co-operation and standardization.

It is useful to distinguish between two types of data maintenance. Creation of backup files is generally short-term, whereas data archiving is long-term. Both should have an expiry (or destruction) date by default, the responsibility for which lies with the originator of the data.

Since the volume of spatial data sets is often very large, there is an important trade-off between data volume and the value of the data sets. This trade-off should be carefully considered when establishing destruction dates. Whenever storage capacity permits, the original data should be saved, and not the products derived from them, since data can be significantly affected by processing.

Updating of spatial data sets can be done on a recurrent basis as well as on an ad hoc basis. In both cases feedback from users is required. Data decay should be taken into consideration including the different rates of change of socio-economic and natural resources data which form part of an integrated system.

Quality and suitability of data are a major concern for the data-base manager. Apart from the deterioration of data quality due to age (i.e. data decay), the possibility exists that data of questionable quality may enter the system. The originators of the data should normally identify these shortcomings, since knowledge of the methods used for data collection and the personnel involved is important information in the assessment of data quality. Whenever possible statistical measures of data quality should be included in the data files. In any event, early recognition of data impurity is vital since unreliable data can "contaminate" products that incorporate good data. Feedback from frequent users, who perhaps can be identified by the system itself, speeds up the identification of suspect data. Complaints about the suitability of the data (e.g. the way data are referenced) may also be obtained by canvassing principal users. In turn, these users should be notified of all measures taken to remedy such problems.

Historic archiving of data is a separate function in which the rarity (uniqueness and value) of the data should be considered. The responsibilities and expiry dates employed should be established as part of a policy of historical records in general and are not discussed further here.

5.2 FINANCIAL SUPPORT

The way in which the financial support is arranged may determine the fate of a particular information system. The funds available are often a major factor for success, but there are other aspects of financial support of equal importance. Reliable and stable funding enables long-range, well-planned activities. An adequate part of the funding should be foreseen in convertible currency if this is a relevant concern in a given national situation.

Many sources of funding have strings attached such as "for pure science only" or "for applications only, no research" which tend to break up some of the vital linkages within an information system. However, the experience of certain creative funding agencies suggests that through the careful and thoughtful allocation of finance, it is possible to promote effective utilization of information.

5.3 LOGISTIC SUPPORT

Logistic support means the availability of all necessary technical facilities for the operation of the information system (spare parts, repairs, stationery, communication, etc.). Each element of the system should have adequate logistical back-up, supported by appropriate skills and funding. This means, for example, that there should be local support for all components of a computer installation. The maintenance of documentation on information-handling procedures should not be overlooked.

5.4 MANAGEMENT OF CHANGE

Growth or change maintenance prepares for new developments, expected or unexpected. It is a managerial function which is beginning to receive major recognition. Within any information system, provisions should be made to ensure adaptability to new knowledge and new circumstances and to keep abreast of the changing needs of the user. It seems probable that information systems have to be open and flexible enough to gear themselves to changes in knowledge and in the technology of knowledge utilization and communication.

An expansion of the system in terms of hardware, peripherals, software or data holdings, such as the integration of historical information through the use of polygon overlays, may lead to a revision of spatial data collection methods. If such a situation is fully exploited, significant reductions in data collection and data holdings can often be achieved. Similarly, new methods and technologies such as high-resolution data-capture devices, decentralization and implementation of new capacities may lead to dramatic modification or discontinuation of certain system modules.

An important responsibility of management is to ensure adequate educational support, since training will be required for new system capacities and skills (see 2.2.3). Transfer of technology and skills will be necessary if the sphere of applications is extended to include other agencies.

Educational support should provide for the training of the professionals, the maintenance of professional standards and, particularly, the preparation of consumers in the use of the information provided. However, the educational practice often seems to place unbalanced emphasis on basics and operational techniques rather than on insight in system development and information use. The documentation and dissemination of experiences from elsewhere should be adequately supported. The importance of updating and retraining the in-house staff and of informing the consumers should not be neglected, even if recruitment and training of new staff must be given priority at a given time.

GUIDELINE 23: To ensure smooth and a continuous operation of the information system, administrative support structures have to be created in the technical, financial, logistical and managerial areas. These support structures have to be carefully planned.

REFERENCES CITED

1. M.Mosaad Allam, "Digital mapping in a distributed system environment". In: J. Foreman (ed.), AUTO-CARTO 5, American Society of Photogrammetry and Congress on Surveying and Mapping, Falls Church, 1983.

2. See:
   - R.F. Tomlinson (ed.), Geographical data handling, Proceedings of the Unesco/IGU 2nd Symposium on Geographical Information Systems, Ottawa-Ont., 1972 and,
   - O. Salomonsson, "Data gaps in the planning process; an application to environmental planning". In: W.H. de Man and A.D.J. Schaap (eds.), Proceedings of the Workshop on Information Requirements for Development Planning in Developing Countries, ITC, Enschede, 1980.

3. J. Bonnen, "Improving information on agriculture and rural life", American Journal of Agricultural Economics, Dec. 1975.

4. Data for Development International Association (DFD) (ed.), Data and development, DFD, Marseille, 1979.

5. H. Ozbekhan, "Towards a general theory of planning". In: E. Jantsch (ed.), Perspectives of planning, OECD, Paris, 1969. Ozbekhan refers to the three levels as "normative, strategic and operational planning".

6. A.D. de Groot, Methodology, Mouton, The Hague, 1969.

7. I. Masser, "The limits to planning", Town Planning Review, Vol. 51/1, 1980.

8. Useful texts on monitoring and/or evaluation are
   - Caroll H. Weiss, Evaluation research; methods of assessing program effectiveness, Prentice-Hall, New Jersey, 1972.
   - Guido J. Deboeck, Guidelines for monitoring and evaluation of rural development projects, Staff Working Papers Series, World Bank, Washington D.C., 1977.
   - N. Imboden, Managing information for rural development projects, OECD, Paris, 1980.

9. The establishment of such a government data and information plan has been recommended (amongst others) by the ESCAP Working Group on Information Systems in Government in collaboration with Data for Development International Association (Marseille). According to Data for Development, a Government Information Plan is the approach for the integrated planning and development of government information systems. It relies on the concept of Government Data Network, which is made up of the following elements: data flows in public administration; procedures for the collection, organization, processing and use of these data; the institutions involved and the resources available (data professionals, data processing equipment, etc.). See Data for Development Newsletter, no's 11 (1980) and 14 (1981), Data for Development International Association, Marseille.

10. The Urban Data Management Software (UDMS) Package, User's Manual, 2nd Edition. United Nations Centre for Human Settlement (HABITAT), P.O. Box 30030, Nairobi, Kenya, January 1983.
UDMS, designed for use with micro-computers, is a general data-management package intended to meet the data storage, analysis, display and mapping needs of a broad range of users such as physical, social, health, environmental and land-use planners, housing officials, resource managers, regional scientists and other professionals.

11. R.F. Tomlinson, "The handling of data for natural resources development". In: W.H. de Man and A.D.J. Schaap (eds.), Proceedings of the Workshop on Information Requirements for Development Planning in Developing Countries, ITC, Enschede, 1980.

12. - R.F. Tomlinson (ed.), Geographical data handling, proceedings of the Unesco/IGU 2nd Symposium on Geographical Information Systems, Ottawa-Ont., 1972.
- O. Salomonsson, "Geocoding for developing countries; a discussion paper", Data-for-Development Newsletter No.13, Data for-Development International Association, Marseille, 1981.
- Central Bureau of Statistics (Kenya), "Geocoding in data systems for Kenya", Kenya Statistical Digest, Vol. XVIII, Nairobi, 1980.
- B.J. Glick and S.A. Hirsch, "An integrated spatial statistics package for map data analysis". In: J. Foreman (ed.), AUTO-CARTO 5, American Society of Photogrammetry and Congress on Surveying and Mapping, Falls Church, 1983.
- J. Dangermond (et al.), "Software components commonly used in geographic information systems". In: D. Marble (ed.), Kensington Workshop, 1981.

13. J. Philip and C.A. Gersmehl, The graphic display of inventory data. In: J. Foreman (ed.), AUTO-CARTO 5, American Society of Photogrammetry and Congress on Surveying and Mapping, Falls Church, 1983. These authors distinguish two somewhat conflicting purposes for a spatial information system: regional inventory and parcel classification for local use. They illustrate the difference between these two by questions to be asked, such as "how much of resource X is in area Y?" (regional inventory) and "what is the resource in parcel Z?" (parcel classification).

14. - M.D. Gwynne and J.J. Croze, "The concept and practice of ecological monitoring over large areas of land: the Systematic Reconnaissance Flight (SRF)". In: Proceedings of the Ibadan/ Garoua International Symposium on Wildlife management in Savanna Woodland, University of Ibadan, 1975.
- M. Norton-Griffiths, T. Hart and M. Parton, "Sample surveys from light aircraft combining visual observation and very large scale colour photography", Remote Sensing Newsletter No. 82/2, Arizona Remote Sensing Centre, Arizona, 1982.

15. J.L. Agatsiva and H. Mwendwa, Land use mapping of Kenya using remote sensing techniques, K.R.E.M.U. Technical Report Series no. 72, Kenya Rangeland Ecological Monitoring Unit, Nairobi, 1982.

16. H.G. Jerie, J. Kure and H.K. Larsen, "A systems approach to improving geo-information systems", ITC Journal Vol. 1980/4 (ITC Post Congress Seminar - Special Issue), ITC, Enschede, 1980.

17. R.G. Havelock and K.D. Benne, "An exploratory study of knowledge utilization". In: G. Watson (ed.), Concepts for social change, National Training Labs., 1966. Also in: W.G. Bennis et al. (eds.), The planning of change, Holt, Rinehart and Winston, 1976.

ANNEX 1

List of participants in the Experts'Meeting on the Framework/Guidelines for Establishing Information Services held at ITC, Enschede, The Netherlands, 27 February - 2 March 1984

Prof. G. T. AGYEPONG	University of Sokoto Sokoto, NIGERIA
Prof. Stein W. BIE	Norwegian Computing Centre Oslo, NORWAY
Mr. Lloyd G. BRETT	Jamaica National Investment Promotion Limited Kingston, JAMAICA
Mr. Helmut EPP	Kenya Rangeland Ecological Monitoring Unit Nairobi, KENYA
Mr. Fulgentio K.S. SERUNJOGI	Bank of Uganda Kampala, UGANDA
Dr. Kavasseri V. SUNDARAM	Planning Commission, Government of India New Delhi, INDIA
Dr. John B. ROSE	Division of the General Information Programme, Unesco Paris, FRANCE
Mr. Sidney W. WITIUK	Spatially-oriented Referencing Systems Association (SORSA) and Statistics Canada Ottawa-Ont., CANADA

From the ITC staff:

Dr. Fouad A.A.F. AMER

Mr. Jan M.M. van den BROEK

Dr. G.W. Wouter ELBERSEN

Prof. Henk A. LUNING

Prof. Jan J. NOSSIN

Mr. Kurt SIJMONS

Dr. Pavao STEFANOVIC

Mr. Erik van der ZEE

Mr. W.H. Erik de MAN (co-ordinator)

Ms. Nana O. KORANTENG (assistant co-ordinator)

Mr. Ver S. SANTOS (assistant co-ordinator)

Mrs. Fely A. de BOER (secretary)

LIST OF RESOURCE PERSONS CONTRIBUTING TO THE PROJECT

Prof. G.T. Agyepong	University of Sokoto Sokoto, Nigeria
Dr. D.C. Alonzo	Chief Statistical Development Service FAO Rome, Italy
Dr. N.E.O Ambitho	Central Bureau of Statistics Ministry of Finance and Planning Nairobi, Kenya
Mr. D.K. Andere	Director Kenya Rangeland Ecological Monitoring Unit (KREMU) Ministry of Finance and Planning Nairobi, Kenya
Dr. P.L.P. Bartelmus	Environment Statistics Sect. United Nations Statistical Office New York, USA
Prof. S.W. Bie	Norwegian Computing Centre Oslo, Norway
Mr. L.G. Brett	Jamaica National Investment Promotion Limited Kingston, Jamaica
Dr. B. Chattopadhyay	Director Centre for Regional, Ecological and Science Studies In Development Alternatives (CRESSIDA) Calcutta, India
Dr. J.C. Coiner	Special Adviser Data Management UN Centre for Human Settlements (HABITAT) Nairobi, Kenya
Dr. D. Cusack (†03.06.1984)	Agro Climate Development Progr. Sierra Blanca Associates Coulder, Colorado, USA
Dr. J.C. Davis	Kansas Geological Survey University of Kansas Lawrence, Kansas, USA
Mr. Y. Deniel	Director THERMODATA Martin d'Heres, France

Prof. J.B. Dent	Dept. Farm Management Lincoln College Canterbury, New Zealand
Prof. H.D. Dias	Human Settlements Div. Asian Institute of Technology Bangkok, Thailand
Dr. I.M. El Azab	Director General of Agric. Stat. Central Agency for Statistics Cairo, Egypt
Mr. H. Epp	Kenya Rangeland Ecological Monitoring Unit Nairobi, Kenya
Mr. M. Frere	Plant Production and Protection Div. FAO Rome, Italy
Mr. A.M. Friend	Office of the Senior Adviser on Integration Statistics Canada Ottawa, Canada
Mr. E. Gura	Directorate-General Inform. Market and Innovation Commission of the European Communities Luxembourg
Dr. J. Hock	Assessment and Information Services Nat. Oceanic and Atmospheric Adm. NOAA Washington, DC., USA
Mr. M. Karaba	National Environment Secretariat Ministry of Environment and Natural Resources Nairobi, Kenya
Mr. P.F. Krumpe	Office of US Foreign Disaster Assistance US-AID Washington, D.C., USA
Mr. A. Laihonen	Central Statistical Office Helsinki, Finland
Mr. R. Leblond	International Development Research Centre Ottawa, Canada
Mr. P.C.T. Mayeye	Bureau of Statistics Ministry of Planning and Economic Affairs Dar es Salaam, Tanzania

Mr. D.B.J. Mutoro	Central Bureau of Statistics Ministry of Finance and Planning Nairobi, Kenya
Mr. J.N. Muttunga	National Environment Secretariat Nairobi, Kenya
Mr. M.J. Njenga	National Environment Secretariat Nairobi, Kenya
Prof. M.S. Patel	Dep. of Mathematics University of Nairobi Nairobi, Kenya
Dr. D. Peden	Kenya Rangeland Ecological Monitoring Unit (KREMU) Ministry of Finance and Planning Nairobi, Kenya
Dr. A.K.M.G. Rabbani	Director-General Central Bureau of Statistics Secretary Statistics Division Ministry of Planning Dacca, Bangladesh
Prof. D.J. Rapport	Office of the Senior Adviser on Integration Statistics Canada Ottawa, Canada
Dr. W.O. Roberts	Former Director Food, Climate and the World's Future Program Aspen Institute for Humanistic Studies Boulder, Colorado, USA
Mr. D.A. Robinson	Rural Physical Planning Unit Ministry of Agriculture Mandeville, Jamaica
Mr. D.C. Rodriguez	Consejo Superior de Planificacion Economica Tegucigalpa D.C., Honduras
Mr. J. Salmona	Director-General Data-for-Development Intern. Association Marseille, France
Mr. R. Sarroca Varela	Oficina de Planification Agricola Ministerio de Agricultura Santiago, Chile

Dr. W. Scott	UN-Research Institute for Social Development Geneva, Switzerland
Mr. F.K.S. Serunjogi	Bank of Uganda Kampala Uganda
Mr. P. Singh	Special Adviser on Statistics CFTC Commonwealth Secretariat London, UK
Mr. L.H. Smith	Ministry of Agriculture Christ Church, Barbados
Dr. Soemantri S.	Director Centre for Environmental Research Gadjah Mada University Yogyakarta,Indonesia
Mr. N.D. Strommen	World Agricultural Outlook Board US Depatment of Agriculture Washington, D.C., USA
Dr. K.V. Sundaram	Planning Commission Government of India New Delhi, India
Dr. G. Thimaiah	Institute for Social and Economic Change Bangalore, India
Mr. S. Tilimo	Land Use Planning Department Addis Abeba, Ethiopia
Mr. J. Torres Martinez	Chief Department of Planification National System of Scient. & Technical Information Havana, Cuba
Mr. H.M. Turcios	Direccion Planification Sectorial Recursos Naturales Tegucigalpa D.C., Honduras
Mr. J.T.M. Uiso	Central Bureau of Statistics Ministry of Planning and Economic Affairs Dar es Salaam, Tanzania

From the ITC, Enschede, The Netherlands:

Dr. F.A.A.F. Amer

Mr. J.M.M. van den Broek

Mr. C.A. de Bruijn

Dr. G.W.W. Elbersen

Dr. H.G.J. Huizing

Prof. H.G. Jerie

Mr. J. Kure

Prof. H.A. Luning

Prof. J.J. Nossin

Mr. V.F.L. Pollé

Mr. A.D.J. Schaap

Mr. K. Sijmons

Dr. P. Stefanovic

Mr. J.H. de Vos t.N.C.

Mr. E. van der Zee

ANNEX 2

SELECTIVE LIST OF EXISTING GEOGRAPHIC INFORMATION SYSTEMS

The following selective list of existing geographic information systems serves primarily the purpose of illustration. The source of this selective list is mainly the 'IGU-Software Inventory', which was made available to the ITC-Workshop on Information Requirements for Development Planning in Developing Countries, Enschede, The Netherlands, 1980, through the kindness and co-operation of the Chairman of the Commission on Geographical Data Sensing and Processing of the International Geographical Union.

CGIS - Canada Geographic Information System.

Canada Land Data System Division, Lands Directorate, Environment Canada, Ottawa Ont., Canada.

CGIS is perhaps the largest and technically most sophisticated system in existence; most of its data i.e. land suitability for uses such as agriculture, forestry, recreation and wildlife habitat was acquired through the Canada Land Inventory (CLI) mapping programme; data from other sources include census data and administrative, watershed and shoreline boundaries.

The approach used in CGIS is, first to transform map data into a structured, non-graphic format amenable to computer processing, and, secondly, to provide a set of procedures that can operate efficiently on the stored data. The creation of the data bank involves the processing of source data by a uniform set of operations. The subsequent manipulation and retrieval of the data involve sets of processes that can vary for each user and each use. The sequence and type of operations may change with each request, for example, overlay, dissolve, or merge. A "monitor" system facilitates and co-ordinates the operations needed to manipulate the data bank, and has also a special compiler that allows a number of commands for graphic data handling to be expressed in an English-like language. The output from requests can be in tabular or plotted-line form or both.

GBF/DIME - Geographic Base File/Dual Independent Map Encoding System (USA).

US Bureau of the Census, Geography Division and Census Use Study, Washington, D.C., USA.

The GBF/DIME System consists of the GBF/DIME-File, a topographic map showing network structure (as a street-network), and a computer program to create and maintain the file. The GBF/DIME-File itself is a computerized representation of the topographic map. (The US Bureau of the Census uses its Metropolitan Map Series as topographic maps.)

The DIME-File is composed of segment records. A segment is defined as a length of a street or other feature between two distinct vertices or nodes. When these segments are connected they create enclosed areas (blocks). DIME (Dual Independent Map Encoding) refers to the fact that the computer file which represents the topographic map is created from two independent sets of identifiers for each segment: (1) the node points for each intersection, and (2) the enclosed areas seperated by each segment.

There are two types of DIME topological edits. One "chains" the string of segments that bound a block and is therefore known as the block chain edit. The other "chains" the blocks surrounding a node and is known as the node chain edit. The block chain edit is performed first and is the most important edit as it detects the great majority of structural defects in the coded file.

Two classes of computer programs are available. The first is concerned with actually creating GBF/DIME-Files and with correcting and updating files. The second type deals with using GBF/DIME-Files as data input for analyzing networks and for allocation purposes.

LUNR - Land Use and Natural Resources Inventory (USA).

Cornell University in collaboration with Albany Office of Planning Services.
Centre for Aerial Photographic Studies, Cornell University, Ithaca, New York, USA.

LUNR is a computerized record of an aerial survey inventory conducted of New York's land use and natural resources. The data, 85% of which was collected through the interpretation of 1:24,000 scale photography, is referenced to a 1 km by 1 km grid-cell.

The LUNR Inventory classification system was devised after consulting with a variety of potential clients including representatives from various state agencies and planners from county and larger areal levels. It consists of ten major categories: agriculture, forest land, water resources, residential, commercial and industrial, outdoor recreation, extractive industry, public and semi-public transportation, and non-productive land. These categories are further subdivided into more detailed areal or point sub-categories. There are about 50 areal and about 70 point sub-categories. Four categories of supplementary data are then added, namely on soils, based on the USDA soil map, on bedrock type, based on the USGS geological map, on the economic viability of farm areas and finally on landforms and depth to bedrock.

MAGI - Maryland Automated Geographic Information System (USA).

Environmental Systems Research Institute, Redlands, California, USA.

The MAGI system has as inputs maps depicting natural resources (soils, vegetation, mineral resources, etc.) and cultural geography (land use, transportation, etc.). These are converted into digital data on a 2000 ft by 2000 ft (610 m) grid cell. Users have six basic models at their disposal for the production of either listings or grid cell overlay displays of combinations of the variables. These are a mining/extraction model; a productive agriculture model; urban models involving an urban centres model, an intensive residential model, a low intensity residential model and an industrial location model; a conservation model; a forestry model; a fishery model. In addition, sub-models have also been developed for public recreation areas, accessibility, natural processes and hazards, agricultural encroachment, tourism and wildlife habitats.

MAP/MODEL - Map/Model Planning Information System (USA).

Bureau of Governmental Research and Services,
University of Oregon, Eugene, Oregon, USA.

The MAP/MODEL system is a series of concepts, procedures and computer programs designed to facilitate the handling and use of (typically) mapped information. The primary spatial units used by the system are polygons, lines or points, which describe the explicit boundaries or locations of data. Digital representations

or 'models' are established and maintained for each type of data such as land use, streets and roads, soils, census tracts, watersheds, zoning, ownership, etc. This concept of data structure allows for substantial flexibility in geocoding. Co-ordinates are used to reference the data. Data in any one 'model' or between two or more 'models' can be quantified, combined, modified, ordered and displayed.

MLMIS - Minesota Land Management Information System (USA).

Minesota State Planning Agency, Land Management Information Centre, St. Paul, Minnesota, USA.

MLMIS is a centralized data bank and analysis system for Minesota's natural resources, developed with the idea to standardize and to some degree centralize at state level the collection and storage of data to be used by the various state and local government agencies. The 40 acre (16 ha) 'parcel' of the U.S. Public Land Survey was chosen as the basic cell size for the system because historically it is the basis for allocating land in the state. Also, the size of the 'parcel' is convenient for computer mapping and analysis of data covering relatively large areas. In areas where land ownership has been fragmented from this basic size, enough information exists to aggregate land back to the original 'parcel'.

The main information types recorded in each 'parcel'are the township, minor civil divisions as defined by the Bureau of the Census, public ownership, type of acquisition of State or Country-owned land, highest recommended use of State or County-owned lands, recommended disposition of State or County-owned lands, management unit status of State or County-owned lands, land use, forest cover, water orientation, highway orientation, soil landscape unit, geomorphic regions. In addition to these statewide variables, some areas of the State have additional information on file.

NIDAS - National Integrated Data System (Malaysia).

Centre for Policy Research at Universiti Sains Malaysia, Penang, Malaysia.

The purpose of NIDAS is to improve the administrative machinery for planning and implementation. The basic files are the land data bank, the census and population register, and the building/property

register. These basic registers are integrated by a geographic base file that allows the display of the spatial attributes of the data elements which the registers contain. In addition to the basic files, the NIDAS data base contains five secondary files, namely a health file, an education and manpower file, an industry/business file, a land utilization file, and a development projects file.

Geocodes for the geographic base file exist in different forms depending not only on the type of geographic data (point data, line data, areal data) but also on the different levels of organization. The co-ordinate referencing system is the most basic, which provides the necessary flexibility for retrieval and integration and computer mapping. Basic geocoding is accomplished by digitizing structures and land parcels and by establishing an address directory. The address directory associates individual land parcels with a specific location. Data elements within the various files may be related by matching addresses or geoco-ordinates. This ability to link files by location is the key to relating one data element of the data base to another.
Since the data base, geocoding and analytical capabilities are still in development, NIDAS is in a constant state of change.

NIMS - A Planning Information System (Sweden).

Nordplan and NIMS-Gruppen, Sundbyberg (POB 133), Sweden.

NIMS has been developed to assist local authority planners in solving recurrent and everyday problem situations. The data in NIMS are derived from both central and local registers. The central registers cover personal data (age, sex, civil status, nationality), real estate data (information from property taxation), and workplace data (main fields of activity and numbers employed), whereas the local registers include roads and highways, traffic accidents, and air raid shelters.

A geographical network is primarily used in NIMS as a means to describe spatial characteristics of objects. The local authority's highway, street, cycle and footpath network is split up into smaller units, known as segments. With segments, flexible areas can be delimited. In addition to segment co-ordinates, NIMS can also store co-ordinates for points and polygons.

The (spatial) interrelationship between the elements of the central registers is facilitated by registers with specific functions such

as the integration registers, which account for the spatial interrelationships between address, property, segment, and area; co-ordinate registers which specify location for properties, segments, and areas; and topological registers defining the spatial interrelationships between segments.

The NIMS-80 system is a general machine-independent information system using relational data base techniques. NIMS-80 has subsystems for various municipal routines such as property registration, fire inspection, building maintenance, electricity and water billing, garbage collection, road maintenance, day care and school administration, etc. Each subsystem is built upon a prototype, a set of user functions, the actual design of which might vary from one municipality to another. When a subsystem is implemented, the user specifications are used to adapt the prototype to the specific local requirements. The structure of NIMS-80 allows this to be done in a very flexible and time saving way by using system parameters. Programming is reduced to a minimum during the adaptation process. The same methods allow the user to change specifications as experience with the system is gained.

ORRMIS - Oak Ridge Regional Modeling Information System (USA).

Regional Urban Studies Dept., Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA.

ORRMIS was developed to provide data management capability for analysis models which forecast the spatial distribution and ecological effects of socio-economic, land use, socio-political activities taking place within a geographical region. Subsystems within ORRMIS are those for geographical data, model communication, remote sensing, topographical analysis, census data and user language.

ANNEX 3

GLOSSARY OF TERMS USED IN THIS REPORT

ACCURACY - the extent of freedom from error (compare with precision).

BASIC SPATIAL UNIT - the smallest spatial reference object for a geographic information system. Examples are: topographic areas, grid-cells, lines, points, addresses. Usually, spatial data is collected for these 'basic spatial units'.

BASIC (SPATIAL) DATA SET - When two spatial data sets are integrated (or linked to each other), each having a different geocoding scheme, a basic spatial data set determines the geocoding scheme of the resulting (integrated) spatial data set in that the other (related) spatial data set are linked to the basic spatial data set according to the geocoding scheme of the latter.

CENTRALIZED INFORMATION SYSTEM - see: modular design.

DECENTRALIZED INFORMATION SYSTEM - see: modular design.

DISTRIBUTED INFORMATION SYSTEM - see: modular design.

GENERALIZATION - selection and simplified representation of detail as appropriate to scale and/or purpose of a map.

GEOCODING - assigning a spatial reference (or identifier)to spatial phenomena (objects). The method and rules according to which geocoding will be done is called the "geocoding scheme". A distinction can be made between direct- and indirect-geocoding. Synonym: geographic referencing.

GEOGRAPHIC BASE FILE - map and code topographic structures/ networks.

GEOGRAPHIC INFORMATION SYSTEM - a means that serves as a provider of spatial data and/or information.
Synonyms: geo-information system, spatial information system.

GRID-CELL - smallest region formed by a (regular or irregular) arbitrary grid; one of the commonly used geocoding schemes.

HIERARCHICAL INFORMATION SYSTEM - see: modular design.

MAP (-LIKE) - spatial data set, representing a region of the earth; normally to scale and usually on a flat medium. A 'map' conforms to conventions in selection and representation (e.g. topographic map, soil map), whereas 'map-like' products are less stringent in this respect (e.g. annotated aerial photographs, satellite imagery).

MODULAR DESIGN - design of an information system as composed of several sub-systems with standardized connections. If the particular composition reflects some hierarcy between the modules, the system is called 'hierarchical'; if the modules are placed in geographically different locations, the system is called 'distributed'; if the modules are placed at lower-level decision-making centres, the system is called 'decentralized'.

MOSAIC - an assembly of aerial photographs, or other remote sensing imagery or maps matched along the edges to provide a continuous representation of part of the earth's surface.

ORIENTATION - inherent to the kind of 'mission' the information system is expected to fulfil. Two different orientations are distinguished: general purpose supply orientation and specialized or function orientation.

POINT - single (x,y) co-ordinate pair.

POLYGON - a closed region bounded by straight line segments.

PRECISION - statistical measure of repeatability; usually expressed as a variance or a standard deviation (compare with accuracy).

RASTER - arbitrarily choosen system of adjacent regular (generally rectangular or square) area units.

RELATED (SPATIAL) DATA SET - see: basic (spatial) data set.

REMOTE SENSING - the collection of data by non-contact measurement methods/techniques.

RESOLUTION - measure of the ability to distinguish between two closely adjacent objects.

SEGMENT - line defined by two terminal points.